Capacity-based random codes cannot achieve strong secrecy over symmetric wiretap channels

Laura Luzzi, Matthieu Bloch

Capacity-based random codes cannot achieve strong secrecy over symmetric wiretap channels

Laura Luzzi
Supélec, 91192 Gif-sur-Yvette, France
laura.luzzi@supelec.fr

Matthieu R. Bloch
School of ECE, Georgia Institute of Technology,
Atlanta, GA
GT-CNRS UMI 2958, Metz, France
matthieu.bloch@ece.gatech.edu

ABSTRACT

In this paper, we investigate the limitations of *capacity-based* random code constructions for the wiretap channel, i.e., constructions that associate to each confidential message a subcode whose rate approaches the capacity of the eavesdropper's channel.
Generalizing a previous result for binary symmetric channels, we show that random capacity-based codes do not achieve the strong secrecy capacity over the symmetric discrete memoryless channels they were designed for. However, we also show that these codes can achieve strong secrecy rates provided they are used over degraded wiretap channels.

1. INTRODUCTION

In most communication schemes, coding at the physical layer is merely performed to ensure reliability; however, seminal results obtained by Wyner and Csiszár & Körner for the wiretap channel [1, 2] have shown the existence of coding schemes that can simultaneously guarantee reliability and secrecy against passive eavesdroppers. In addition, the secrecy of such schemes is measured quantitatively in terms of statistical independence. Specifically, if the random variable M represents the transmitted message, X^n represents the encoded message and Z^n represents the observation of an eavesdropper, perfect secrecy is achieved if M and Z^n are independent or, equivalently, $\mathbb{I}(\mathsf{M};\mathsf{Z}^n)=0$. Exact statistical independence is unfortunately too stringent to be amenable to further analysis; therefore, as originally suggested by Wyner, it is convenient to relax the constraint and to require *asymptotic* statistical independence in the limit of large encoding length n. Two measures of asymptotic statistical independence have been commonly used:

- weak secrecy, which requires $\lim_{n\to\infty}\frac{1}{n}\mathbb{I}(\mathsf{M};\mathsf{Z}^n)=0$;
- strong secrecy, which requires $\lim_{n\to\infty}\mathbb{I}(\mathsf{M};\mathsf{Z}^n)=0$;

It has been shown that the maximum rate of secure and reliable communication over a wiretap channel is the same irrespective of the metric used [3]; however, this does not imply that a specific code achieving weak secrecy achieves strong secrecy. With the exception of [4, 5], most code constructions based on polar codes [6, 7, 8, 9] or LDPC codes [10] proposed thus far have been proved to achieve weak secrecy only and whether they achieve strong secrecy remains an open question. A closer look at the latter constructions reveals that secrecy is obtained by associating to each confidential message a subcode whose rate approaches the capacity of the eavesdropper's channel; hence we call such codes *capacity-based* wiretap codes.

In this paper, we highlight the potential limitations of capacity-based wiretap codes by showing that random capacity-based wiretap codes that achieve the weak secrecy capacity for symmetric discrete memoryless channels (DMCs) cannot achieve the strong secrecy capacity. This result is a generalization of that obtained in [11] for binary symmetric channels. The proof follows the same reasoning as [11], but we reproduce it here in full for the sake of completeness.

The rest of the paper is organized as follows. Section 2 introduces the notation and channel model considered in the paper. Section 3 summarizes known techniques to show the achievability of secure rates and highlights how capacity-based codes guarantee secrecy. Section 4 forms the core of the paper and proves that random capacity-based wiretap codes cannot achieve the strong secrecy capacity of symmetric DMCs. The special case where the channel of the legitimate receiver is error-free is also examined. Moreover, it is shown that capacity-based codes can achieve strong secrecy rates provided they are used over degraded wiretap channels. Section 5 offers some concluding remarks and perspectives.

2. PRELIMINARIES

2.1 Notation

We briefly detail the notation used in the paper. The notation log always stands for the logarithm in base 2. For random variables X, Y, $\mathbb{H}(\mathsf{X})$ denotes the entropy of X, and $\mathbb{I}(\mathsf{X};\mathsf{Y})$ the mutual information between X and Y. For probability distributions p, q, $\mathbb{D}(p\|q)$ and $\mathbb{V}(p,q)$ denote respectively the Kullback-Leibler divergence and L^1 variational distance of p and q.

Given a proposition $\mathcal{P}$, we define

$$\mathbf{1}_{\{\mathcal{P}\}} = \begin{cases} 1 & \text{if } \mathcal{P} \text{ is true,} \\ 0 & \text{if } \mathcal{P} \text{ is false.} \end{cases}$$

2.2 Wiretap channel

We consider the wiretap channel $\mathrm{WT}(W_b, W_e)$ illustrated in Figure 1, which consists of two symmetric DMCs: the channel from $\mathcal{X}$ to $\mathcal{Y}$ of the legitimate receiver with transition probabilities $W_b = p_{\mathsf{Y}|\mathsf{X}}$ and capacity C_b, and the channel from $\mathcal{X}$ to $\mathcal{Z}$ of the eavesdropper, with transition probabilities $W_e = p_{\mathsf{Z}|\mathsf{X}}$ and capacity C_e. The alphabets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are assumed to be finite. The notion of symmetric channel used in the paper and its properties are detailed in Section 2.3.

Definition (**Wiretap code**). *A $(2^{nR}, 2^{nR'}, n)$ wiretap code $\mathcal{C}_n$ consists of a message set $\mathcal{M}_n = [\![1, 2^{nR}]\!]$, an auxiliary message set $\mathcal{M}'_n = [\![1, 2^{nR'}]\!]$, an encoding function $f_n : \mathcal{M}_n \times \mathcal{M}'_n \to \mathcal{X}^n$, and a decoding function (for the legitimate receiver) $g_n : \mathcal{Y}^n \to \mathcal{M}_n \times \mathcal{M}'_n$. M denotes the confidential message, and M' the auxiliary message, X^n is the transmitted codeword and Y^n, Z^n are the corresponding outputs of the channels W_b and W_e. $(\hat{\mathsf{M}}, \hat{\mathsf{M}}') = g_n(\mathsf{Y}^n)$ is the estimate of the legitimate receiver.*

The messages M and M' are assumed to be uniformly distributed in their respective sets. The reliability of a wiretap code is measured in terms of the average probability of error

$$P_e(\mathcal{C}_n) \triangleq \mathbb{P}\left\{(\mathsf{M}, \mathsf{M}') \neq (\hat{\mathsf{M}}, \hat{\mathsf{M}}')|\mathcal{C}_n\right\}.$$

while its secrecy is measured in terms of the information leaked to the eavesdropper

$$\mathbb{L}(\mathcal{C}_n) \triangleq \mathbb{I}(\mathsf{M}; \mathsf{Z}^n|\mathcal{C}_n).$$

Definition (**Capacity-based and resolvability-based codes**). *A sequence of wiretap codes $\{\mathcal{C}_n\}_{n\geq 1}$ is called* capacity-based *if*

$$R' = C_e - \varepsilon_n, \quad \lim_{n\to\infty} \varepsilon_n = 0,$$

and if $\forall n > 0$ there exists a decoding function $h_n : \mathcal{Z}^n \times \mathcal{M}_n \to \mathcal{M}'_n$. We denote by $\tilde{\mathsf{M}}' = h_n(\mathsf{Z}^n, \mathsf{M})$ the corresponding estimate.
If on the contrary $R' > C_e$ for large n, we say that the sequence of codes is resolvability-based.

The reliability of a capacity-based wiretap code is measured with the modified average probability of error

$$P_e^*(\mathcal{C}_n) = \mathbb{P}\left\{(\hat{\mathsf{M}}, \hat{\mathsf{M}}') \neq (\mathsf{M}, \mathsf{M}') \text{ or } \tilde{\mathsf{M}}' \neq \mathsf{M}'|\mathcal{C}_n\right\}.$$

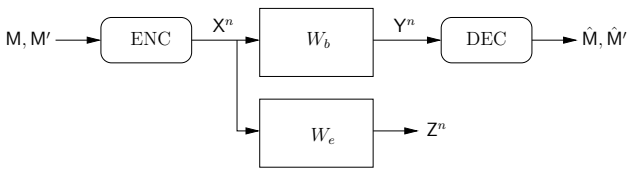

Figure 1: Wiretap channel model.

Definition (**Achievable secrecy rates**). *A rate R is an achievable* weak *secrecy rate if there exists a sequence of wiretap codes $\{\mathcal{C}_n\}_{n\geq 1}$ of rate R such that*

$$\lim_{n\to\infty} P_e(\mathcal{C}_n) = 0 \quad \lim_{n\to\infty} \frac{1}{n}\mathbb{L}(\mathcal{C}_n) = 0.$$

Similarly, a rate R is an achievable strong *secrecy rate if there exists a sequence of wiretap codes $\{\mathcal{C}_n\}_{n\geq 1}$ of rate R such that*

$$\lim_{n\to\infty} P_e(\mathcal{C}_n) = 0 \quad \lim_{n\to\infty} \mathbb{L}(\mathcal{C}_n) = 0.$$

The rate R is an achievable (strong or weak) secrecy rate with capacity-based wiretap codes if the same conditions hold upon replacing P_e by P_e^*. The weak (resp. strong) secrecy capacity C_s is the supremum of achievable weak (resp. strong) secrecy rates.

2.3 Properties of symmetric channels

In this paper we focus on the case where W_b and W_e are symmetric DMCs. Several notions of channel symmetry have been proposed in the literature [12]. For instance, the following definition was given by Cover and Thomas [13].
Recall that the transition matrix of a channel W from $\mathcal{X} = \{x_1, \ldots, x_{|\mathcal{X}|}\}$ to $\mathcal{Z} = \{z_1, \ldots, z_{|\mathcal{Z}|}\}$ is

$$\mathbf{M} = (m_{ij}) = (W(z_j|x_i)).$$

Definition (**CT-symmetric**). *A DMC W is CT-symmetric if every row of the transition matrix $\mathbf{M}$ is a permutation of every other row, and all the column sums are equal.*

This condition is too restrictive for us, for it does not include the binary erasure channel, one of the few examples of channels for which explicit wiretap codes achieving both weak and strong secrecy have been constructed [4, 5, 6, 14]. We will thus adopt the following alternative condition proposed by Gallager [15]:

Definition (**G-symmetric**). *A DMC W with input $\mathcal{X}$ and output $\mathcal{Z}$ is G-symmetric if $\mathcal{Z}$ can be partitioned into subsets in such a way that, for every subset, the corresponding submatrix of transition probabilities has the property that each row is a permutation of each other row and each column is a permutation of each other column. That is, there exists a partition $\mathcal{Z} = \mathcal{Z}_1 \cup \cdots \cup \mathcal{Z}_r$ into disjoint sets such that:*

- *$\forall a, \bar{a} \in \mathcal{X}$, there exists a permutation $\pi_{a\bar{a}} : \mathcal{Z} \to \mathcal{Z}$ such that $\forall k \in [\![1, r]\!]$, $\pi_{a\bar{a}}(\mathcal{Z}_k) = \mathcal{Z}_k$, and*

$$\forall z \in \mathcal{Z}, \quad W(z|a) = W(\pi_{a\bar{a}}(z)|\bar{a}), \tag{1}$$

- *$\forall k \in [\![1, r]\!]$, $\forall b, \bar{b} \in \mathcal{Z}_k$, there exists a permutation $\pi_{b\bar{b}} : \mathcal{X} \to \mathcal{X}$ such that*

$$\forall x \in \mathcal{X}, \quad W(b|x) = W(\bar{b}|\pi_{b\bar{b}}(x)). \tag{2}$$

Note that G-symmetry does not imply CT-symmetry, and vice-versa [12].

Theorem (**Gallager**). *For a G-symmetric DMC with input $\mathcal{X}$ and output $\mathcal{Y}$, the input distribution that achieves capacity is the uniform distribution on $\mathcal{X}$.*

Remark 1. *Even though the output distribution q_{Z} corresponding to the uniform input distribution q_{X} is not necessarily uniform for G-symmetric channels, it turns out that it is locally uniform on each set $\mathcal{Z}_k$, $k = 1, \ldots, r$ of the partition [15]:*

$$\forall z, \bar{z} \in \mathcal{Z}_k, \quad q_{\mathsf{Z}}(z) = q_{\mathsf{Z}}(\bar{z}).$$

This property follows easily from the fact that the columns are permutations of each other:

$$q_{\mathsf{Z}}(z) = \sum_{a \in \mathcal{X}} q_{\mathsf{X}}(a) W(z|a) = \sum_{a \in \mathcal{X}} \frac{1}{|\mathcal{X}|} W(z|a) =$$
$$= \sum_{a \in \mathcal{X}} \frac{1}{|\mathcal{X}|} W(\bar{z}|\pi_{z\bar{z}}(a)) = \sum_{a' \in \mathcal{X}} \frac{1}{|\mathcal{X}|} W(\bar{z}|a') = q_{\mathsf{Z}}(\bar{z}).$$

Remark 2. *For a G-symmetric channel W from $\mathcal{X}$ to $\mathcal{Z}$, if the input X is uniformly distributed on $\mathcal{X}$,*

$$\forall x, \bar{x} \in \mathcal{X}, \quad \mathbb{D}(p_{\mathsf{Z}|\mathsf{X}=x} \| q_{\mathsf{Z}}) = \mathbb{D}(p_{\mathsf{Z}|\mathsf{X}=\bar{x}} \| q_{\mathsf{Z}}).$$

Consequently, the capacity $C = \mathbb{I}(\mathsf{X};\mathsf{Z}) = \mathbb{D}(p_{\mathsf{Z}|\mathsf{X}=x} \| q_{\mathsf{Z}}) \ \ \forall x \in \mathcal{X}$.

Proof. We have

$$\mathbb{D}(p_{\mathsf{Z}|\mathsf{X}=x} \| q_{\mathsf{Z}}) = \sum_{z \in \mathcal{Z}} W(\pi_{x\bar{x}}(z)|\bar{x}) \log \frac{W(\pi_{x\bar{x}}(z)|\bar{x})}{q_{\mathsf{Z}}(\pi_{x\bar{x}}(z))} =$$
$$= \sum_{z' \in \mathcal{Z}} W(z'|\bar{x}) \log \frac{W(z'|\bar{x})}{q_{\mathsf{Z}}(z')} = \mathbb{D}(p_{\mathsf{Z}|\mathsf{X}=\bar{x}} \| q_{\mathsf{Z}}). \qquad \square$$

We can write

$$\mathbb{I}(\mathsf{X};\mathsf{Z}) = \sum_{x \in \mathcal{X}} q_{\mathsf{X}}(x) \sum_{z \in \mathcal{Z}} W(z|x) \log \frac{W(z|x)}{q_{\mathsf{Z}}(z)} =$$
$$= \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} \mathbb{D}(p_{\mathsf{Z}|\mathsf{X}=x} \| q_{\mathsf{Z}}).$$

3. ACHIEVING SECRECY: CAPACITY VS. RESOLVABILITY

In this section, we prove that random capacity-based wiretap codes achieve weak secrecy capacity and show that random resolvability-based wiretap codes achieve strong secrecy. Strictly speaking, these proofs have already appeared in various forms [1, 16, 17, 18], and we reproduce them to highlight the fundamental differences between the two constructions.

In both situations, we start with a random capacity-based code $\mathcal{C}_n$, whose $2^{n(R+R')}$ codeword symbols are generated independently according to a distribution q_{X} on $\mathcal{X}$. The codewords are labeled $c(m, m')$ with $m \in [\![1, 2^{nR}]\!]$ and $m' \in [\![1, 2^{nR'}]\!]$. Let q_{Z} be the output distribution of the eavesdropper's channel corresponding to the input q_{X}, defined as

$$\forall z \in \mathcal{Z}, \quad q_{\mathsf{Z}}(z) = \sum_{x \in \mathcal{X}} W_e(z|x) q_{\mathsf{X}}(x),$$

and q_{Z^n} the product of n i.i.d. copies of this output distribution:

$$q_{\mathsf{Z}^n}(z^n) = \prod_{i=1}^{n} q_{\mathsf{Z}}(z_i).$$

Lemma 1 (Secrecy from capacity). *Let C_n denote the random variable representing the randomly generated code. Let $\epsilon_n > 0$ be such that $\lim_{n\to\infty} \epsilon_n = 0$ but $\lim_{n\to\infty} \sqrt{n}\epsilon_n = \infty$. Then, for n sufficiently large, there exists $\alpha > \beta > 0$ such that*

$$\left\{ \begin{array}{l} R + R' < \mathbb{I}(\mathsf{X};\mathsf{Y}) \\ R' = \mathbb{I}(\mathsf{X};\mathsf{Z}) - \epsilon_n \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \mathbb{E}_{\mathsf{C}_n}\{P_e^*(\mathsf{C}_n)\} \leq 2^{-\alpha n} \\ \mathbb{E}_{\mathsf{C}_n}\left\{\frac{1}{n}\mathbb{L}(\mathsf{C}_n)\right\} \leq \epsilon_n + \frac{1}{n}2^{-\beta n}. \end{array} \right.$$

Proof. The fact that the conditions $R + R' < \mathbb{I}(\mathsf{X};\mathsf{Y})$ and $R' < \mathbb{I}(\mathsf{X};\mathsf{Z})$ imply $\mathbb{E}_{\mathsf{C}_n}\{P_e^*(\mathsf{C}_n)\} \leq 2^{-\alpha n}$ for some $\alpha > 0$ follows from standard large deviation techniques, see for instance [19]. Next, notice that

$$\begin{aligned}
\mathbb{E}_{\mathsf{C}_n}\{\mathbb{L}(\mathsf{C}_n)\} &= \mathbb{E}_{\mathsf{C}_n}\{\mathbb{I}(\mathsf{M};\mathsf{Z}^n|\mathsf{C}_n)\} \\
&= \mathbb{E}_{\mathsf{C}_n}\{\mathbb{I}(\mathsf{M}\mathsf{X}^n;\mathsf{Z}^n|\mathsf{C}_n) - \mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n|\mathsf{M})\} \\
&= \mathbb{E}_{\mathsf{C}_n}\{\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n|\mathsf{C}_n) + \mathbb{I}(\mathsf{M};\mathsf{Z}^n|\mathsf{X}^n\mathsf{C}_n) \\
&\qquad -\mathbb{H}(\mathsf{X}^n|\mathsf{M}) + \mathbb{H}(\mathsf{X}^n|\mathsf{M}\mathsf{Z}^n)\} \\
&\overset{(a)}{\leq} \mathbb{E}_{\mathsf{C}_n}\{\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n) - \mathbb{H}(\mathsf{X}^n|\mathsf{M}) + \mathbb{H}(\mathsf{X}^n|\mathsf{M}\mathsf{Z}^n)\} \\
&\overset{(b)}{\leq} \mathbb{E}_{\mathsf{C}_n}\{n\mathbb{I}(\mathsf{X};\mathsf{Z}) - nR' + P_e^*(\mathsf{C}_n)nR\} \\
&\leq n\epsilon_n + nR\mathbb{E}_{\mathsf{C}_n}\{P_e^*(\mathsf{C}_n)\} \\
&\leq n\epsilon_n + nR2^{-\alpha n} \\
&\overset{(c)}{\leq} n\epsilon_n + 2^{-\beta n}.
\end{aligned}$$

Here (a) follows from the fact that $\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n|\mathsf{C}_n) < \mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n)$, and that $\mathbb{I}(\mathsf{M};\mathsf{Z}^n|\mathsf{X}^n,\mathsf{C}_n) = 0$ because $\mathsf{M} \to \mathsf{X}^n \to \mathsf{Z}^n$ is a Markov chain; (b) follows from Fano's inequality and (c) holds for some $\beta \in (0, \alpha)$ and n sufficiently large. $\square$

Note that the speed at which $\mathbb{E}_{\mathsf{C}_n}\{\mathbb{L}(\mathsf{C}_n)\}$ decays is tightly related to the speed at which one can approach the capacity of the eavesdropper's channel $\mathbb{I}(\mathsf{X};\mathsf{Z})$. Unfortunately, large deviation techniques cannot circumvent the condition $\lim_{n\to\infty} \sqrt{n}\epsilon_n = \infty$.

Instead of using capacity-based wiretap codes, one may use resolvability-based wiretap codes, in which case the analysis of secrecy relies on the following lemma [20].

Lemma 2 (Cloud mixing). *If the random codebook size is $2^{n\bar{R}}$ with $\bar{R} > \mathbb{I}(\mathsf{X};\mathsf{Z})$, then $\exists \beta > 0$ such that*

$$\forall n, \quad \mathbb{E}_{\mathsf{C}_n}\left[\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{C}_n}, q_{\mathsf{Z}^n})\right] \leq e^{-\beta n},$$

where the expectation is computed over the random code ensemble.

Lemma 3 (Secrecy from resolvability). *Let C_n denote the random variable representing the randomly generated code. Then, for n sufficiently large, there exists $\alpha > \gamma > 0$ such that*

$$R + R' < \mathbb{I}(\mathsf{X};\mathsf{Y}) \Rightarrow \mathbb{E}_{\mathsf{C}_n}\{P_e(\mathsf{C}_n)\} \leq 2^{-\alpha n}$$
$$R' > \mathbb{I}(\mathsf{X};\mathsf{Z}) \Rightarrow \mathbb{E}_{\mathsf{C}_n}\{\mathbb{L}(\mathsf{C}_n)\} \leq 2^{-\gamma n}.$$

Proof. The first part follows from the same arguments as in Lemma 1.

The second part follows by noting that

$$\begin{aligned}\mathbb{E}_{\mathsf{C}_n}\{\mathbb{V}(p_{\mathsf{M}\mathsf{Z}^n|\mathsf{C}_n}, p_{\mathsf{M}}\times p_{\mathsf{Z}^n|\mathsf{C}_n})\}&\\ &= \mathbb{E}_{\mathsf{C}_n\mathsf{M}}\{\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}\mathsf{C}_n}, p_{\mathsf{Z}^n|\mathsf{C}_n})\}\\ &\le \mathbb{E}_{\mathsf{C}_n\mathsf{M}}\{\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}\mathsf{C}_n}, q_{\mathsf{Z}^n}) + \mathbb{V}(q_{\mathsf{Z}^n}, p_{\mathsf{Z}^n|\mathsf{C}_n})\}\\ &\le 2\mathbb{E}_{\mathsf{C}_n\mathsf{M}}\{\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}\mathsf{C}_n}, q_{\mathsf{Z}^n})\}\\ &\overset{(a)}{=} \mathbb{E}_{\mathsf{C}_n}\{\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=1\mathsf{C}_n}, q_{\mathsf{Z}^n})\}\\ &\overset{(b)}{\le} 2^{-\beta n}\end{aligned}$$

where (a) follows by symmetry of the random code construction and (b) follows from Lemma 2. Then, Lemma 1 in [21] guarantees that there exists $\gamma > 0$ such that

$$\mathbb{E}_{\mathsf{C}_n}\{\mathbb{L}(\mathsf{C}_n)\} \le 2^{-\gamma n}. \qquad \square$$

Note that the condition $R' > \mathbb{I}(\mathsf{X};\mathsf{Z})$ allows us to establish strong secrecy directly. Naturally, one can wonder whether our inability to prove strong secrecy in Lemma 1 is fundamentally linked to the use of capacity-based codes, or whether this is simply a limitation of the proof. In the following section, we show that random capacity-based wiretap codes cannot achieve the strong secrecy capacity, which suggests that such constructions are less powerful than resolvability-based ones.

4. CAPACITY-BASED RANDOM CODES DO NOT ACHIEVE THE STRONG SECRECY CAPACITY

We consider a random capacity-based code $\mathcal{C}_n$ whose codewords are generated independently according to the *uniform* distribution q_{X} on $\mathcal{X}$. To simplify notation, we denote by $p_{\mathsf{X}^n} = p_{\mathsf{X}^n|\mathcal{C}_n}$ the uniform distribution on the codewords of $\mathcal{C}_n$, and $p_{\mathsf{Z}^n} = p_{\mathsf{Z}^n|\mathcal{C}_n}$ the corresponding output distribution of the eavesdropper's channel. Note that, in general, the components of p_{Z^n} are not i.i.d.

The following Proposition generalizes the result of Lemma 3 in [11] obtained for the case of binary symmetric channels:

Proposition 1. *Let $\{\mathcal{C}_n\}_{\ge 1}$ be a sequence of $(2^{nR}, 2^{nR'}, n)$ capacity-based codes for the wiretap channel $WT(W_b, W_e)$ obtained by generating codeword symbols independently according to the uniform distribution q_{X} on $\mathcal{X}$, and such that $R + R' < C_b$, the channel capacity of the legitimate receiver. Then there exist $\alpha' > 0$, $\eta > 0$ such that $\forall \kappa > 0$,*

$$\mathbb{P}_{\mathsf{C}_n}\{\mathbb{L}(\mathcal{C}_n) \ge \eta - \kappa - f_\kappa(n)\} \ge 1 - 2^{-\alpha' n}.$$

for some function $f_\kappa : \mathbb{N} \to \mathbb{R}^+$ with $\lim_{n\to\infty} f_\kappa(n) = 0$.

Proof. By definition, $\mathbb{L}(\mathcal{C}_n) = \mathbb{I}(\mathsf{M};\mathsf{Z}^n) = \mathbb{D}(p_{\mathsf{M}\mathsf{Z}^n}\|p_{\mathsf{M}}\times p_{\mathsf{Z}^n})$. Watanabe *et al.* [22] (Theorem 6 and proof of Theorem 7) showed that $\forall b > 0$,

$$\mathcal{E}(\mathcal{C}_n) + \mathbb{V}(p_{\mathsf{M}\mathsf{Z}^n}, p_{\mathsf{M}}\times p_{\mathsf{Z}^n}) \ge 1 - (2^{-b\sqrt{n}+1} + \mathbb{P}(\mathcal{A}_n)),$$

where $\mathcal{E}(\mathcal{C}_n) = \mathbb{P}\left\{\tilde{\mathsf{M}}' \ne \mathsf{M}'\right\}$, and

$$\begin{aligned}\mathcal{A}_n &= \left\{(x^n, z^n) : \frac{2^{-b\sqrt{n}}}{|\mathcal{M}_n|} < p_{\mathsf{X}^n|\mathsf{Z}^n}(x^n|z^n) \le \frac{2^{b\sqrt{n}}}{|\mathcal{M}_n|}\right\} =\\ &= \left\{(x^n, z^n) : \frac{2^{-b\sqrt{n}}}{|\mathcal{M}_n|} < \frac{p_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n)p_{\mathsf{X}^n}(x^n)}{p_{\mathsf{Z}^n}(z^n)} \le \frac{2^{b\sqrt{n}}}{|\mathcal{M}_n|}\right\}\end{aligned}$$

Clearly $P_e^*(\mathcal{C}_n) \ge \mathcal{E}(\mathcal{C}_n)$. From Pinsker's inequality, we get

$$\begin{aligned}&\mathbb{V}(p_{\mathsf{M}\mathsf{Z}^n}, p_{\mathsf{M}}\times p_{\mathsf{Z}^n}) \le \sqrt{2\ln 2\, \mathbb{D}(p_{\mathsf{M}\mathsf{Z}^n}\|p_{\mathsf{M}}\times p_{\mathsf{Z}^n})} =\\ &= \sqrt{2\ln 2\, \mathbb{L}(\mathcal{C}_n)},\end{aligned}$$

therefore

$$P_e^*(\mathcal{C}_n) + \sqrt{2\ln 2\, \mathbb{L}(\mathcal{C}_n)} \ge 1 - \left(2^{-b\sqrt{n}+1} + \mathbb{P}(\mathcal{A}_n)\right). \quad (3)$$

Since X^n is uniform on the code, $p_{\mathsf{X}^n}(\mathsf{X}^n) = \frac{1}{|\mathcal{C}_n|} = \frac{1}{2^{n(R+R')}} = \frac{1}{|\mathcal{M}_n|2^{n(C_e-\varepsilon_n)}}$. Therefore we can write

$$\begin{aligned}&\mathbb{P}(\mathcal{A}_n) = \mathbb{P}(\mathcal{A}_n^+) - \mathbb{P}(\mathcal{A}_n^-), \quad \text{where}\\ &\mathcal{A}_n^+ = \left\{\log\frac{p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{Z}^n|\mathsf{X}^n)}{p_{\mathsf{Z}^n}(\mathsf{Z}^n)} \le b\sqrt{n} + n(C_e - \varepsilon_n)\right\},\\ &\mathcal{A}_n^- = \left\{\log\frac{p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{Z}^n|\mathsf{X}^n)}{p_{\mathsf{Z}^n}(\mathsf{Z}^n)} \le -b\sqrt{n} + n(C_e - \varepsilon_n)\right\}. \quad (4)\end{aligned}$$

Estimate of $\mathbb{P}(\mathcal{A}_n^+)$. The set $\mathcal{A}_n^+$ can be rewritten as

$$\left\{\log\frac{p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{Z}^n|\mathsf{X}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} - \log\frac{p_{\mathsf{Z}^n}(\mathsf{Z}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} \le b\sqrt{n} + n(C_e - \varepsilon_n)\right\}$$

Let

$$\begin{aligned}\mathcal{B}_n &= \left\{\log\frac{p_{\mathsf{Z}^n}(\mathsf{Z}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} < b\sqrt{n}\right\},\\ \mathcal{A}_n' &= \left\{\log\frac{p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{Z}^n|\mathsf{X}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} \le 2b\sqrt{n} + n(C_e - \varepsilon_n)\right\}.\end{aligned}$$

By the law of total probability,

$$\begin{aligned}&\mathbb{P}(\mathcal{A}_n^+) = \mathbb{P}(\mathcal{A}_n^+|\mathcal{B}_n)\mathbb{P}(\mathcal{B}_n) + \mathbb{P}(\mathcal{A}_n^+|\mathcal{B}_n^c)\mathbb{P}(\mathcal{B}_n^c) \le\\ &\le \mathbb{P}(\mathcal{A}_n^+ \cap \mathcal{B}_n) + \mathbb{P}(\mathcal{B}_n^c) \le \mathbb{P}(\mathcal{A}_n') + \mathbb{P}(\mathcal{B}_n^c)\end{aligned}$$

since $\mathcal{A}_n^+ \cap \mathcal{B}_n \subset \mathcal{A}_n'$. We have

$$\begin{aligned}&\mathbb{P}(\mathcal{B}_n^c) = \mathbb{P}\left\{\log\frac{p_{\mathsf{Z}^n}(\mathsf{Z}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} \ge b\sqrt{n}\right\} =\\ &= \frac{1}{b\sqrt{n}}\sum_{z^n\in\mathcal{Z}^n} b\sqrt{n}\, p_{\mathsf{Z}^n}(z^n)\mathbf{1}_{\left\{\log\frac{p_{\mathsf{Z}^n}(z^n)}{q_{\mathsf{Z}^n}(z^n)} \ge b\sqrt{n}\right\}} \le\\ &\le \frac{1}{b\sqrt{n}}\sum_{z^n\in\mathcal{Z}^n} p_{\mathsf{Z}^n}(z^n)\log\frac{p_{\mathsf{Z}^n}(z^n)}{q_{\mathsf{Z}^n}(z^n)}\mathbf{1}_{\left\{\log\frac{p_{\mathsf{Z}^n}(z^n)}{q_{\mathsf{Z}^n}(z^n)} \ge b\sqrt{n}\right\}} \le\\ &\le \frac{1}{b\sqrt{n}}\mathbb{D}(p_{\mathsf{Z}^n}\|q_{\mathsf{Z}^n}).\end{aligned}$$

We can estimate the divergence as follows:

$$\begin{aligned}&\mathbb{D}(p_{\mathsf{Z}^n}\|q_{\mathsf{Z}^n}) = \sum_{z^n\in\mathcal{Z}^n} p_{\mathsf{Z}^n}(z^n)\log\frac{p_{\mathsf{Z}^n}(z^n)}{q_{\mathsf{Z}^n}(z^n)} =\\ &= -\mathbb{H}(p_{\mathsf{Z}^n}) - \sum_{z^n\in\mathcal{Z}^n} p_{\mathsf{Z}^n}(z^n)\log q_{\mathsf{Z}^n}(z^n) =\\ &= \mathbb{H}(q_{\mathsf{Z}^n}) - \mathbb{H}(p_{\mathsf{Z}^n}) + \sum_{z^n\in\mathcal{Z}^n}(q_{\mathsf{Z}^n}(z^n) - p_{\mathsf{Z}^n}(z^n))\log q_{\mathsf{Z}^n}(z^n).\end{aligned}$$

Recall that the entropy is continuous with respect to the variational distance (Lemma 2.7 in [23]): if two probability distributions p,q on $\mathcal{X}$ satisfy $\mathbb{V}(p,q) \le \frac{1}{2}$, then

$$|\mathbb{H}(p) - \mathbb{H}(q)| \le \mathbb{V}(p,q)\log\left(\frac{|\mathcal{X}|}{\mathbb{V}(p,q)}\right).$$

Therefore

$$|\mathbb{H}(p_{\mathsf{Z}^n}) - \mathbb{H}(q_{\mathsf{Z}^n})| \leq \mathbb{V}(p_{\mathsf{Z}^n}, q_{\mathsf{Z}^n}) \log\left(\frac{|\mathcal{Z}|^n}{\mathbb{V}(p_{\mathsf{Z}^n}, q_{\mathsf{Z}^n})}\right).$$

Since the rate of the random code is $R + R' = R + C_e - \varepsilon_n > C_e = \mathbb{I}(\mathsf{X};\mathsf{Z})$ for n large enough, Lemma 2 holds. Markov's inequality implies that for $\beta < \alpha$,

$$\mathbb{P}\left(\mathbb{V}(p_{\mathsf{Z}^n}, q_{\mathsf{Z}^n}) \geq 2^{-\beta n}\right) \leq 2^{-\alpha n + \beta n} \leq 2^{-\alpha' n}$$

for some $\alpha' > 0$, for n large enough. So with probability greater than $1 - 2^{-\alpha' n}$, the first term can be bounded by

$$|\mathbb{H}(q_{\mathsf{Z}^n}) - \mathbb{H}(p_{\mathsf{Z}^n})| \leq 2^{-\beta n}(n \log|\mathcal{Z}| + \beta n). \tag{5}$$

The second term in the expression of the divergence is in turn bounded by

$$\begin{aligned}
&- \sum_{z^n \in \mathcal{Z}^n} (p_{\mathsf{Z}^n}(z^n) - q_{\mathsf{Z}^n}(z^n)) \log q_{\mathsf{Z}^n}(z^n) \leq \\
&\leq \mathbb{V}(p_{\mathsf{Z}^n}, q_{\mathsf{Z}^n}) \max_{z^n \in \mathcal{Z}^n} (-\log q_{\mathsf{Z}^n}(z^n)) = n\mathbb{V}(p_{\mathsf{Z}^n}, q_{\mathsf{Z}^n}) \log\frac{1}{\mu_{\mathsf{Z}}} \leq \\
&\leq \log\frac{1}{\mu_{\mathsf{Z}}} n 2^{-\beta n} \qquad (6)
\end{aligned}$$

where $\mu_{\mathsf{Z}} = \min_{z \in \mathrm{Supp}(q_{\mathsf{Z}})} q_{\mathsf{Z}}(z)$. Thus, from (5) and (6) we conclude that

$$\mathbb{P}(\mathcal{B}_n^c) \leq \frac{2^{-\beta n}}{b\sqrt{n}}\left(n \log|\mathcal{Z}| + \beta n + n \log\frac{1}{\mu_{\mathsf{Z}}}\right) \leq Cn2^{-\beta n}$$

for some $C > 0$. We still need to show that

$$\begin{aligned}
&\mathbb{P}(\mathcal{A}_n') = \mathbb{P}\left\{\sum_{i=1}^n \log\frac{W_e(\mathsf{Z}_i|\mathsf{X}_i)}{q_{\mathsf{Z}}(\mathsf{Z}_i)} \leq \varphi(n)\right\} = \\
&= \sum_{\substack{x^n \in \mathcal{X}^n \\ z^n \in \mathcal{Z}^n}} p_{\mathsf{X}^n}(x^n) p_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n) \mathbf{1}_{\left\{\sum_{i=1}^n \log\frac{W_e(z_i|x_i)}{q_{\mathsf{Z}}(z_i)} \leq \varphi(n)\right\}}
\end{aligned}$$

also vanishes, where $\varphi(n) = 2b\sqrt{n} + n(C_e - \varepsilon_n)$. For a fixed realization $\mathcal{C}_n = \{c(1), \ldots, c(|\mathcal{C}_n|)\}$ of the code, this can be written as the weighted sum over the codewords:

$$\sum_{j=1}^{|\mathcal{C}_n|} \frac{1}{|\mathcal{C}_n|} \sum_{z^n \in \mathcal{Z}^n} \prod_{i=1}^n W_e(z_i|c(j)_i) \mathbf{1}_{\left\{\sum_{i=1}^n \log\frac{W_e(z_i|c(j)_i)}{q_{\mathsf{Z}}(z_i)} \leq \varphi(n)\right\}}.$$

Let $x, \bar{x} \in \mathcal{X}$: from the property (1) in the definition of G-symmetric channel and from Remark 1, we have

$$\frac{W_e(z|x)}{q_{\mathsf{Z}}(z)} = \frac{W_e(\pi_{x\bar{x}}(z)|\bar{x})}{q_{\mathsf{Z}}(z)} = \frac{W_e(\pi_{x\bar{x}}(z)|\bar{x})}{q_{\mathsf{Z}}(\pi_{x\bar{x}}(z))}.$$

Since the $\pi_{x\bar{x}} : \mathcal{Z} \to \mathcal{Z}$ are permutations, all the terms in the sum over the codewords coincide and are equal to

$$\begin{aligned}
&\sum_{z^n \in \mathcal{Z}^n} \left(\prod_{i=1}^n W_e(z_i|\bar{x})\right) \mathbf{1}_{\left\{\sum_{i=1}^n \log\frac{W_e(z_i|\bar{x})}{q_{\mathsf{Z}}(z_i)} \leq \varphi(n)\right\}} = \\
&= \mathbb{P}\left\{\sum_{i=1}^n \log\frac{W_e(\widetilde{\mathsf{Z}}_i|\bar{x})}{q_{\mathsf{Z}}(\widetilde{\mathsf{Z}}_i)} \leq \varphi(n)\right\}.
\end{aligned}$$

where $\widetilde{\mathsf{Z}}_i$ is the random variable corresponding to the output of the eavesdropper's channel for a fixed input equal to $\bar{x}$. Since the channel is memoryless, the $\widetilde{\mathsf{Z}}_i$, $i \in [\![1, n]\!]$, are independent and identically distributed and the random variables

$$\mathsf{E}_i = f(\widetilde{\mathsf{Z}}_i) = \log\frac{W_e(\widetilde{\mathsf{Z}}_i|\bar{x})}{q_{\mathsf{Z}}(\widetilde{\mathsf{Z}}_i)}$$

are also i.i.d. with common pdf p_{E}. Moreover, Remark 2 implies that $\mathbb{E}\{p_{\mathsf{E}}\} = C_e$, the capacity of the eavesdropper's channel.
Denote by σ^2 the variance of E, and by ρ its third moment. Observe that $\sigma > 0$ and $\rho < \infty$. Then the Berry-Esseen theorem implies that $\exists c > 0$ such that $\forall x$,

$$\mathbb{P}\left\{\frac{\sum_{i=1}^n (\mathsf{E}_i - C_e)}{\sigma\sqrt{n}} \leq x\right\} \leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{x^2}{2}} dx + \frac{c\rho}{\sqrt{n}\sigma^3}.$$

If we choose $x = \frac{2b}{\sigma} - \frac{\sqrt{n}\varepsilon_n}{\sigma}$, we get

$$\begin{aligned}
\mathbb{P}(\mathcal{A}_n') &= \mathbb{P}\left\{\sum_{i=1}^n \mathsf{E}_i \leq 2b\sqrt{n} + n(C_e - \varepsilon_n)\right\} \leq \\
&\leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{2b}{\sigma} - \frac{\sqrt{n}\varepsilon_n}{\sigma}} e^{-\frac{x^2}{2}} dx + \frac{c\rho}{\sqrt{n}\sigma^3}. \qquad (7)
\end{aligned}$$

and

$$\mathbb{P}(\mathcal{A}^+) \leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{2b}{\sigma} - \frac{\sqrt{n}\varepsilon_n}{\sigma}} e^{-\frac{x^2}{2}} dx + \frac{c\rho}{\sqrt{n}\sigma^3} + C\sqrt{n}2^{-\beta n}.$$

Estimate of $\mathbb{P}(\mathcal{A}^-)$. We estimate the measure of $\mathcal{A}^-$ in a similar manner to $\mathcal{A}^+$. First we rewrite $\mathcal{A}^-$ as

$$\left\{\log\frac{p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{Z}^n|\mathsf{X}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} - \log\frac{p_{\mathsf{Z}^n}(\mathsf{Z}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} \leq -b\sqrt{n} + n(C_e - \varepsilon_n)\right\}$$

Let

$$\begin{aligned}
\mathcal{A}_n'' &= \left\{\log\frac{p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{Z}^n|\mathsf{X}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} \leq -2b\sqrt{n} + n(C_e - \varepsilon_n)\right\}, \\
\mathcal{D}_n &= \left\{\log\frac{p_{\mathsf{Z}^n}(\mathsf{Z}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} \geq -b\sqrt{n}\right\}
\end{aligned}$$

By the law of total probability,

$$\begin{aligned}
\mathbb{P}(\mathcal{A}^-) &= \mathbb{P}(\mathcal{A}_n^+|\mathcal{D}_n)\mathbb{P}(\mathcal{D}_n) + \mathbb{P}(\mathcal{A}_n^+|\mathcal{D}_n^c)\mathbb{P}(\mathcal{D}_n^c) \geq \\
&\geq \mathbb{P}(\mathcal{A}_n^+|\mathcal{D}_n)\mathbb{P}(\mathcal{D}_n) \geq \mathbb{P}(\mathcal{A}_n''|\mathcal{D}_n)\mathbb{P}(\mathcal{D}_n)
\end{aligned}$$

since $\mathcal{A}_n'' \cap \mathcal{D}_n \subset \mathcal{A}_n^+ \cap \mathcal{D}_n$.
We have $\mathbb{P}(\mathcal{D}_n) \geq 1 - 2^{-b\sqrt{n}}$, since

$$\begin{aligned}
\mathbb{P}(\mathcal{D}_n^c) &= \mathbb{P}\left\{\log\frac{p_{\mathsf{Z}^n}(\mathsf{Z}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} < -b\sqrt{n}\right\} = \\
&= \sum_{z^n \in \mathcal{Z}^n} p_{\mathsf{Z}^n}(z^n) \mathbf{1}_{\left\{\log\frac{p_{\mathsf{Z}^n}(z^n)}{q_{\mathsf{Z}^n}(z^n)} < -b\sqrt{n}\right\}} < \\
&\leq \sum_{z^n \in \mathcal{Z}^n} q_{\mathsf{Z}^n}(z^n) 2^{-b\sqrt{n}} \mathbf{1}_{\left\{\log\frac{p_{\mathsf{Z}^n}(z^n)}{q_{\mathsf{Z}^n}(z^n)} < -b\sqrt{n}\right\}} \leq 2^{-b\sqrt{n}}.
\end{aligned}$$

By the Inclusion-Exclusion principle,

$$\begin{aligned}
&\mathbb{P}(\mathcal{A}_n''|\mathcal{D}_n)\mathbb{P}(\mathcal{D}_n) = \mathbb{P}(\mathcal{A}_n'' \cap \mathcal{D}_n) = \\
&= \mathbb{P}(\mathcal{A}_n'') + \mathbb{P}(\mathcal{D}_n) - \mathbb{P}(\mathcal{A}_n'' \cup \mathcal{D}_n) \geq \mathbb{P}(\mathcal{A}_n'') + \mathbb{P}(\mathcal{D}_n) - 1 \geq \\
&\geq \mathbb{P}(\mathcal{A}_n'') - 2^{-b\sqrt{n}}.
\end{aligned}$$

Berry-Esseen's theorem with $x = -\frac{2b}{\sigma} - \frac{\sqrt{n}\varepsilon_n}{\sigma}$ implies that

$$\mathbb{P}(\mathcal{A}_n'') = \mathbb{P}\left\{\log\frac{p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{Z}^n|\mathsf{X}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} \leq -2b\sqrt{n} + n(C_e - \varepsilon_n)\right\} \geq$$

$$\geq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{2b}{\sigma}-\frac{\sqrt{n}\varepsilon_n}{\sigma}} e^{-\frac{x^2}{2}} dx - \frac{c\rho}{\sqrt{n}\sigma^3}. \quad (8)$$

Thus

$$\mathbb{P}(\mathcal{A}^-) \geq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{2b}{\sigma}-\frac{\sqrt{n}\varepsilon_n}{\sigma}} e^{-\frac{x^2}{2}} dx - \frac{c\rho}{\sqrt{n}\sigma^3} - 2^{-b\sqrt{n}}.$$

Estimate of $\mathbb{P}(\mathcal{A}_n)$. Putting together the estimates of $\mathbb{P}(\mathcal{A}^+)$ and $\mathbb{P}(\mathcal{A}^-)$, we can conclude that $\forall b > 0$, with probability greater than $1 - 2^{-\alpha' n}$,

$$\mathbb{P}(\mathcal{A}_n) \leq Cn2^{-\beta n} + 2^{-b\sqrt{n}} + \frac{1}{\sqrt{2\pi}} \int_{-\frac{2b}{\sigma}-\frac{\sqrt{n}\varepsilon_n}{\sigma}}^{\frac{2b}{\sigma}-\frac{\sqrt{n}\varepsilon_n}{\sigma}} e^{-\frac{x^2}{2}} dx +$$
$$+ \frac{2c\rho}{\sqrt{n}\sigma^3} \leq Cn2^{-\beta n} + 2^{-b\sqrt{n}} + \frac{4b}{\sqrt{2\pi}\sigma} + \frac{2c\rho}{\sqrt{n}\sigma^3}$$

for some constant $C > 0$. Then from (3) we get

$$\mathbb{L}(\mathcal{C}_n) \geq \frac{1}{\sqrt{2\ln 2}} \left(1 - f(b,n) - \frac{4b}{\sqrt{2\pi}\sigma} - P_e^*(\mathcal{C}_n)\right),$$

with $\lim_{n\to\infty} f(b,n) = 0$. Lemma 1 implies that with probability tending to 1 exponentially fast, $P_e^*(\mathcal{C}_n) \to 0$. Since b can be arbitrarily small, $\lim_{n\to\infty} \mathbb{L}(\mathcal{C}_n) \geq \frac{1}{\sqrt{2\ln 2}}$. □

4.1 Case of error-free legitimate channel

If the legitimate channel W_b is the identity channel I, the code $\mathcal{U}_n$ consisting of all the possible codewords of length n taken with equal probability always guarantees reliable communication for the legitimate receiver. If each confidential message is associated to a subcode approaching the capacity C_e of the eavesdropper's channel, the sequence $\{\mathcal{U}_n\}_{n\geq 1}$ is a capacity-based code sequence that achieves the weak secrecy capacity. However, a similar result to Proposition 1 still holds in this case. This result has already been proved in [6] with a slightly different approach.

Lemma 4. *Suppose that the sequence of codes* $\{\mathcal{U}_n\}_{n\geq 1}$ *achieves the weak secrecy capacity of the wiretap channel* $WT(I, W_e)$. *Then,* $\exists \eta > 0$ *such that*

$$\lim_{n\to\infty} \mathbb{L}(\mathcal{U}_n) \geq \eta,$$

and so it cannot achieve the strong secrecy capacity.

Proof. The proof of this Lemma is a special case of the proof of Proposition 1. In this case, $p_{\mathsf{X}^n}(\mathsf{X}^n) = q_{\mathsf{X}^n}(\mathsf{X}^n) = \frac{1}{|\mathcal{X}|^n}$ is the uniform distribution. Similarly to equation (4), we can write

$$\mathbb{P}(\mathcal{A}_n) = \mathbb{P}(\mathcal{A}_n^+) - \mathbb{P}(\mathcal{A}_n^-), \quad \text{where}$$
$$\mathcal{A}_n^+ = \left\{ \log \frac{p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{Z}^n|\mathsf{X}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} \leq b\sqrt{n} + n(C_e - \varepsilon_n) \right\},$$
$$\mathcal{A}_n^- = \left\{ \log \frac{p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{Z}^n|\mathsf{X}^n)}{q_{\mathsf{Z}^n}(\mathsf{Z}^n)} \leq -b\sqrt{n} + n(C_e - \varepsilon_n) \right\}.$$

Similarly to equations (7) and (8), after applying Berry-Esseen's theorem we find

$$\mathbb{P}(\mathcal{A}_n^+) \leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{b}{\sigma}-\frac{\sqrt{n}\varepsilon_n}{\sigma}} e^{-\frac{x^2}{2}} dx + \frac{c\rho}{\sqrt{n}\sigma^3},$$

$$\mathbb{P}(\mathcal{A}_n^-) \geq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{b}{\sigma}-\frac{\sqrt{n}\varepsilon_n}{\sigma}} e^{-\frac{x^2}{2}} dx - \frac{c\rho}{\sqrt{n}\sigma^3}.$$

Therefore, $\forall b > 0$,

$$\mathbb{P}(\mathcal{A}_n) \leq \frac{1}{\sqrt{2\pi}} \int_{-\frac{b}{\sigma}-\frac{\sqrt{n}\varepsilon_n}{\sigma}}^{\frac{b}{\sigma}-\frac{\sqrt{n}\varepsilon_n}{\sigma}} e^{-\frac{x^2}{2}} dx + \frac{2c\rho}{\sqrt{n}\sigma^3} \leq$$
$$\leq \frac{2b}{\sqrt{2\pi}\sigma} + \frac{2c\rho}{\sqrt{n}\sigma^3}.$$

The thesis then follows from (3) since $P_e^*(\mathcal{U}_n) \to 0$ (by a similar reasoning to Lemma 1). □

4.2 Achieving strong secrecy by transmitting beyond the eavesdropper's channel capacity

The result of Proposition 1 is rather disappointing, since it shows that capacity-based random codes designed for a symmetric wiretap channel $\mathrm{WT}(W_b, W_e)$ fail to achieve strong secrecy rates over this channel. The good news is that strong secrecy is achievable on all symmetric channels $\mathrm{WT}(W_b, W_e')$ such that the new eavesdropper's channel W_e' has smaller capacity than W_e.

Lemma 5. *Let* $\{\mathcal{C}_n\}_{n\geq 1}$ *be a sequence of random capacity-based codes for the symmetric discrete memoryless wiretap channel* $WT(W_b, W_e)$, *and let* W_e' *be a symmetric DMC with capacity* $C_e' < C_e$. *Then there exists* $\alpha > 0$ *such that with probability greater than* $1 - 2^{-\alpha n}$, $\{\mathcal{C}_n\}_{n\geq 1}$ *achieves strong secrecy rates on* $WT(W_b, W_e')$.

The proof follows the same reasoning as Lemma 3.

5. CONCLUSION AND PERSPECTIVES

In this paper we have generalized a previous result by Bloch [11] for the binary symmetric wiretap channel to all symmetric discrete memoryless wiretap channels, showing that capacity-based code constructions that achieve weak secrecy are suboptimal from the point of view of the strong secrecy metric, and that in particular they cannot achieve the strong secrecy capacity.

Up to now, we have focused on the limitations of capacity-based *random* codes. It is natural to wonder whether *structured* capacity-based codes, such as polar codes for the binary symmetric wiretap channel [6] or the capacity-based LDPC codes for the binary erasure wiretap channel in [14], share the same drawbacks or not. Indeed, in the case where the channel of the legitimate receiver is noiseless, we can already answer affirmatively to this question.

References

[1] A. D. Wyner, "The Wire-Tap Channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1367, October 1975.

[2] I. Csiszár and J. Körner, "Broadcast Channels with Confidential Messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[3] U. M. Maurer and S. Wolf, "Information-Theoretic Key Agreement: From Weak to Strong Secrecy for Free," in *Advances in Cryptology - Eurocrypt 2000*, Lecture Notes in Computer Science. B. Preneel, 2000, p. 351.

[4] A. Subramanian, A. T. Suresh, A. Thangaraj, M. Bloch, and S. McLaughlin, "Strong and Weak Secrecy in Wiretap Channels," in *Proc. of 6th International Symposium on Turbo Codes and Iterative Information Processing*, Brest, France, September 2010, pp. 30 – 34.

[5] A. T. Suresh, A. Subramanian, A. Thangaraj, M. Bloch, and S. McLaughlin, "Strong Secrecy for Erasure Wiretap Channels," in *Proc. IEEE Information Theory Workshop*, Dublin, Ireland, September 2010.

[6] H. Mahdavifar and A. Vardy, "Achieving the Secrecy Capacity of Wiretap Channels Using Polar Codes," in *Proc. of IEEE International Symposium on Information Theory*, Austin, TX, June 2010, pp. 913–917. [Online]. Available: arXiv:1001.0210v1

[7] E. Hof and S. Shamai, "Secrecy-achieving polar-coding," in *Proc. IEEE Information Theory Workshop*, Dublin, Ireland, September 2010, pp. 1–5.

[8] O. O. Koyluoglu and H. E. Gamal, "Polar Coding for Secure Transmission and Key Agreement," 2010, accepted at PIMRC 2010. [Online]. Available: arXiv:1003.1422v1

[9] M. Andersson, V. Rathi, R. Thobaben, J. Kliewer, and M. Skoglund, "Nested Polar Codes for Wiretap and Relay Channels," *IEEE Communications Letters*, vol. 14, no. 4, pp. 752–754, June 2010.

[10] V. Rathi, M. Andersson, R. Thobaben, J. Kliewer, and M. Skoglund, "Two edge type LDPC codes for the wiretap channel," in *Proc. Conf Signals, Systems and Computers Record of the Forty-Third Asilomar Conf*, 2009, pp. 834–838.

[11] M. Bloch, "Achieving secrecy: capacity vs. resolvability," in *Proc. Int. Symp. Inform. Theory (ISIT 2011)*, St. Petersburg, Russia, July-August 2011, to appear.

[12] B. Xie and R. Wesel, "A mutual information invariance approach to symmetry in discrete memoryless channels," in *Information Theory and Applications Workshop*, San Diego, CA, August 2008, pp. 444–448.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 2006.

[14] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of LDPC Codes to the Wiretap Channels," *IEEE Transactions on Information Theory*, vol. 53, no. 8, pp. 2933–2945, Aug. 2007.

[15] R. G. Gallager, *Information Theory and reliable communication*. Wiley, 1968.

[16] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 44–55, 2005.

[17] M. Hayashi, "General Nonasymptotic and Asymptotic Formulas in Channel Resolvability and Identification Capacity and their Application to the Wiretap Channels," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1562–1575, April 2006.

[18] M. Bloch and J. N. Laneman, "On the Secrecy Capacity of Arbitrary Wiretap Channels," in *Proceedings of 46th Allerton Conference on Communication, Control, and Computing*, Monticello, IL, September 2008, pp. 818–825.

[19] G. Kramer, *Topics in Multi-User Information Theory*, ser. Foundations and Trends in Communications and Information Theory. NOW Publishers, 2008, vol. 4, no. 4-5.

[20] P. W. Cuff, "Communication in networks for coordinating behaviour," Ph.D. dissertation, Princeton University, 2009.

[21] I. Csiszár, "Almost Independence and Secrecy Capacity," *Problems of Information Transmission*, vol. 32, no. 1, pp. 40–47, January-March 1996.

[22] S. Watanabe, T. Saitou, R. Matsumoto, and T. Uyematsu, "Strongly secure privacy amplification cannot be obtained by encoder of Slepian-Wolf codes," in *Proc. Int. Symp. Inform. Theory*, Seoul, Korea, July 2009, pp. 1298–1302.

[23] I. Csiszar and J. Korner, *Information Theory: coding theorems for discrete memoryless systems*. Akademiai Kiado, December 1981.